

Conservation Genetics Resources

TECHNICAL NOTE

SNP discovery in the northern dragonhead *Dracocephalum ruyschiana*

Oddmund Kleven, Anders Endrestøl, Marianne Evju, Odd E. Stabbetorp and Kristine B. Westergaard

Oddmund Kleven (✉) – Kristine Bakke Westergaard

Norwegian Institute for Nature Research (NINA), P.O. Box 5685 Torgarden, NO-7485 Trondheim, Norway

Correspondence: Oddmund Kleven, email: oddmund.kleven@nina.no

Anders Endrestøl – Marianne Evju – Odd E. Stabbetorp

Norwegian Institute for Nature Research (NINA), Gaustadalléen 21, NO-0349 Oslo, Norway

Running title: SNP discovery in *Dracocephalum ruyschiana*

Abstract

The northern dragonhead *Dracocephalum ruyschiana* is a plant species experiencing a dramatic population decline that has led to the species being listed on Red Lists for species in many European countries. Here we used restriction-site associated DNA sequencing to isolate and characterize a panel of 96 novel SNP markers from 44 individuals encompassing most of the species range in Norway. The 96 SNPs were adapted for the Fluidigm platform and evaluated by screening another 24 northern dragonheads from a population in southern Norway. The panel of SNP markers developed here are expected to be useful for elucidating genetic diversity and population genetic structure in the northern dragonhead.

Keywords *Dracocephalum ruyschiana* – Genetic diversity – Population genetics – SNP

The northern dragonhead *Dracocephalum ruyschiana* L. is a diploid ($2n=2x=14$), perennial, insect-pollinated herb belonging to the mint family (Lamiaceae) (Lid 2005). It is a Eurasian steppe species with a fragmented distribution, reaching its northwestern limit in Norway, and prefers shallow, calcareous soils in dry meadows and rocky outcrops (Lid 2005). Due to severe reductions in population sizes all over Europe, the northern dragonhead is listed on the Bern Convention Appendix I (<https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/104>) and on many national Red Lists for species. In Norway the species is categorized as vulnerable (VU) on the national Red List for species due to population size reductions and habitat loss (Solstad et al. 2015). To ensure long-term survival of the northern dragonhead in Norway, an action plan has been made for the species (Directorate for Nature Management 2010). The action plan emphasized the knowledge gap concerning genetic diversity and population genetic structure of the species. Hitherto, a lack of available genetic markers has, however, hampered genetic surveys of the northern dragonhead.

The aim of this study was to develop a panel of single-nucleotide polymorphism (SNP) markers to facilitate monitoring of genetic variation as well as studies of population genetic structure and landscape genetic connectivity in the northern dragonhead. We applied a restriction-site associated DNA sequencing method to identify a panel of 96 novel SNP markers from 44 individuals encompassing most of the species range in Norway. To enable rapid and cost-effective genotyping we adapted the SNPs to the Fluidigm system (BioMark - Fluidigm Corporation, San Francisco, USA). The panel of SNPs was evaluated by genotyping another 24 individuals from a population in southern Norway.

Leaves were collected in June and July from 2012 to 2014 and immediately stored in plastic zip-lock bags containing silica-beads, and later ground using tungsten carbide beads and TissueLyser II (Qiagen, Hilden, Germany). DNA was isolated using either the DNeasy plant mini kit (Qiagen) or NucleoSpin plant II extraction kit (Macherey-Nagel, Düren, Germany) following the manufacturers protocols. DNA was eluted in *tris*-EDTA buffer and 48 samples were sent to Ecogenics GmbH (Balgach, Switzerland) for sequencing. In brief, a double-digest restriction-site associated DNA (ddRAD) sequencing approach with EcoRI/MseI was applied. A total of 400ng gDNA per sample was digested and ligated to the respective Illumina adaptors. A small fragment removal step was applied, and the libraries were amplified with Illumina primers containing the respective multiplex identification tags. The tagged libraries were pooled and the size range of 400-500 base-pairs (bp) extracted using gel electrophoresis. The resulting pool was sequenced on a NextSeq chip using the 1×150bp format. The RAD-tags were processed using Stacks (Catchen et al. 2013). Sequence data was obtained for 44 of the 48 individuals, representing 13 geographically separated localities (Figure 1).

SNPs with a minor allele frequency (MAF) less than 0.05 and those with a flanking sequence on each side less than 20 bp were removed. Sequences for the final set of SNPs are provided in the electronic supplementary material (Table S1). Primer design for the Fluidigm SNP type assay was conducted by using the software D3 (<https://d3.fluidigm.com/>). Primer sequences for the final set of SNPs are provided in the electronic supplementary material (Table S1). SNPs were genotyped on a 96.96 Dynamic Array using the Fluidigm EP1 instrument according to the manufacturer's protocol and scored using the Fluidigm SNP genotyping analysis software (<https://www.fluidigm.com/software>).

80

81 Allele frequencies and fixation index was calculated using GenAlEx ver. 6.5 (Peakall and
82 Smouse 2012). Arlequin ver. 3.5.1.2 (Excoffier and Lischer 2010) was used to calculate
83 observed and expected heterozygosities, and to test for deviation from Hardy-Weinberg and
84 linkage equilibrium. A Bonferroni correction for multiple statistical tests (Rice 1989) was
85 applied to linkage disequilibrium p-values.

86

87 One-hundred and forty-two candidate SNPs were tested on the Fluidigm platform. Based on
88 clustering performance and interpretation (data not shown), we selected 96 SNPs. The final
89 set of 96 SNPs was then used to genotype 24 northern dragonheads from a population in
90 southern Norway. Four of the 96 SNPs were monomorphic in this population (Table 1). For
91 the 92 variable SNPs, the mean observed heterozygosity was 0.32 (range 0.04 to 0.67) and
92 mean expected heterozygosity was 0.33 (range 0.04 to 0.51). A single SNP (Dru_34029_70)
93 deviated significantly from Hardy-Weinberg equilibrium. After correcting for multiple tests,
94 significant linkage disequilibrium was detected for two (Dru_30966_28 – Dru_23429_73 and
95 Dru_8642_69 – Dru_21326_30) out of 4560 locus combinations. In conclusion, these novel
96 SNP markers and the Fluidigm SNP-typing assay will be valuable tools in genetic
97 conservation of the northern dragonhead.

98

99 **Acknowledgements**

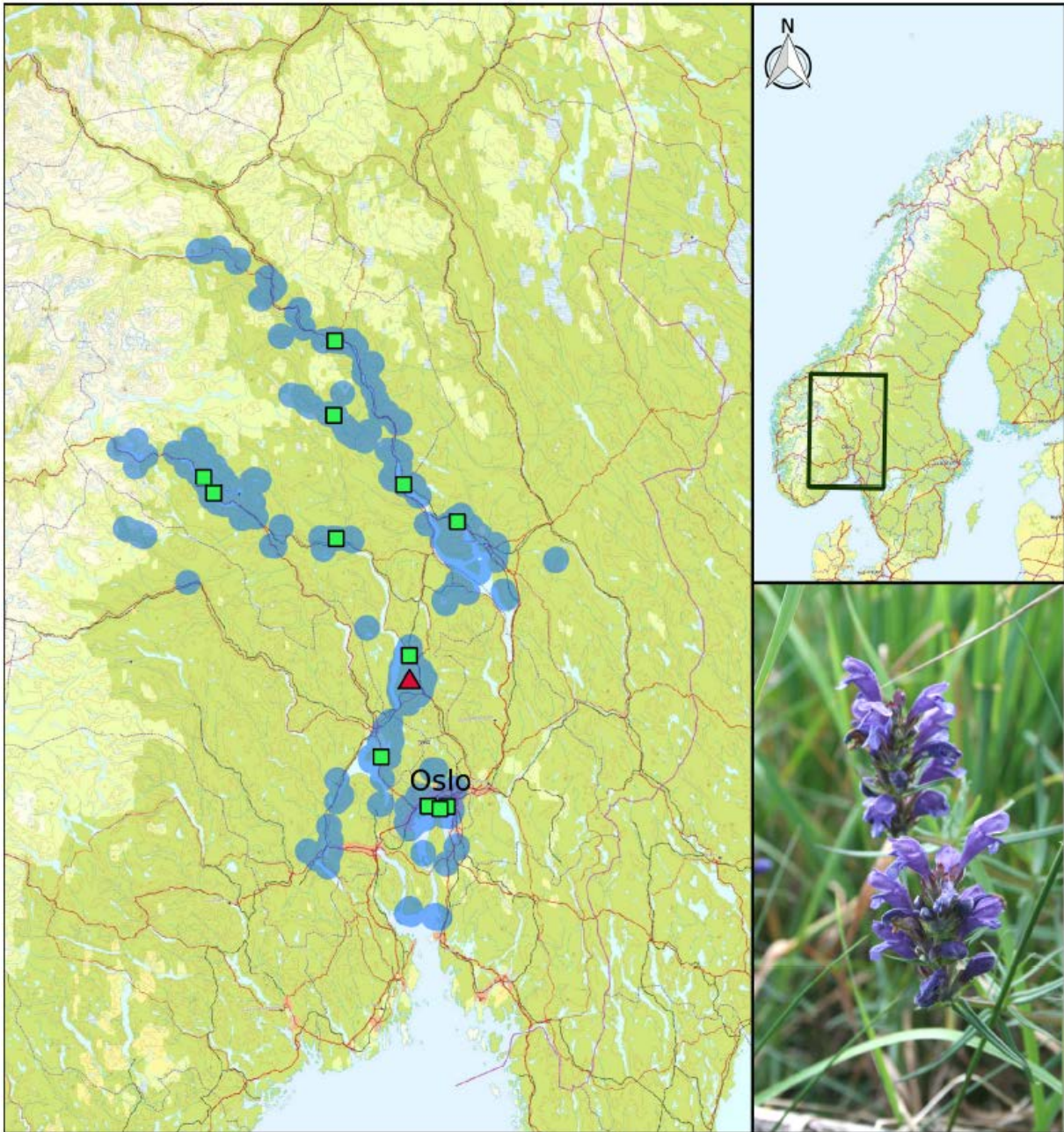
100 We are grateful to Marie K. Brandrud for field assistance and Line B. Eriksen for laboratory
101 assistance. This study was conducted as part of the research projects “Management of
102 biodiversity and ecosystem services in spatially structured landscapes” (208434/F40) and

“Development and implementation of genetic methods“ and financed by the Research Council of Norway and the Norwegian Institute for Nature Research. Permissions to sample leaves were provided by the Directorate for Nature Management (2012 2009/3206 NAK-NK-STB and 2013 2013/3149 NAK-NK-STB) and the Norwegian Environment Agency (2014/4159).

References

- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22 (11):3124-3140. doi:10.1111/mec.12354
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* 10 (3):564-567. doi:10.1111/j.1755-0998.2010.02847.x
- Lid J (2005) Norsk flora. In: Elven R (ed). 7 edn. Det Norske Samlaget, Oslo
- Management DfN (2010) Action plan for the northern dragonhead *Dracocephalum ruyschiana* and dragonhead sap beetle *Meligethes norvegicus* Trondheim. DN-report 2010-5
- Peakall R, Smouse PE (2012) GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics* 28 (19):2537-2539. doi:10.1093/bioinformatics/bts460
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43 (1):223-225
- Solstad H, Elven R, Mjelde M, Pedersen O, Alsos IG, Stabbetorp OE, Gaarder G (2015) Vascular plants. Pteridophyta, Pinophyta, Magnoliophyta. In: Henriksen S, Hilmo O (eds) Norwegian Red List for Species. Norwegian Biodiversity Information Centre, Trondheim, Norway

Figure 1 Geographical distribution of northern dragonhead sampling localities. Squares indicate sampling localities of samples used for sequencing, triangle indicate sampling area for samples used to validate the 96-SNP typing assay. Circles indicates the species' distribution in Norway after 1950 based on data obtained from Species Map Service 1.6 (<https://artskart1.artsdatabanken.no>).



134 **Table 1** Characterization of 96 SNP markers from the northern dragonhead *Dracocephalum*
 135 *ruyschiana*.

Locus ID	SNP identity	Frequency allele 1	Frequency allele 2	MAF	H _O	H _E	P _{HWE}	F
Dru_292_65	A/G	0.21	0.79	0.21	0.42	0.34	0.539	-0.26
Dru_3751_66	C/T	0.50	0.50	0.50	0.42	0.51	0.433	0.17
Dru_4213_33	A/G	0.21	0.79	0.21	0.42	0.34	0.540	-0.26
Dru_4575_29	A/G	0.04	0.96	0.04	0.08	0.08	1.000	-0.04
Dru_6348_29	C/G	0.00	1.00	0.00	0.00	*	*	*
Dru_6458_94	G/T	0.23	0.77	0.23	0.38	0.36	1.000	-0.06
Dru_6533_35	G/T	0.81	0.19	0.19	0.21	0.31	0.153	0.32
Dru_6809_69	C/G	0.13	0.88	0.13	0.25	0.22	1.000	-0.14
Dru_6838_83	A/T	0.19	0.81	0.19	0.21	0.31	0.153	0.32
Dru_6968_77	C/T	0.33	0.67	0.33	0.42	0.45	1.000	0.06
Dru_7068_55	A/G	0.17	0.83	0.17	0.25	0.28	0.502	0.10
Dru_7417_55	A/G	0.33	0.67	0.33	0.42	0.45	1.000	0.06
Dru_7669_91	C/T	0.25	0.75	0.25	0.42	0.38	1.000	-0.11
Dru_7680_55	G/T	0.71	0.29	0.29	0.58	0.42	0.127	-0.41
Dru_8004_59	C/T	0.25	0.75	0.25	0.33	0.38	0.596	0.11
Dru_8642_69	A/C	0.04	0.96	0.04	0.08	0.08	1.000	-0.04
Dru_8798_39	C/T	0.04	0.96	0.04	0.08	0.08	1.000	-0.04
Dru_8930_81	C/T	0.77	0.23	0.23	0.29	0.36	0.556	0.17
Dru_9153_36	C/T	0.25	0.75	0.25	0.42	0.38	1.000	-0.11
Dru_9158_34	C/T	0.65	0.35	0.35	0.46	0.47	1.000	0.00
Dru_9551_25	G/T	0.46	0.54	0.46	0.42	0.51	0.434	0.16
Dru_10600_83	C/T	0.23	0.77	0.23	0.29	0.36	0.556	0.17

Dru_10722_92	G/T	0.60	0.40	0.40	0.54	0.49	0.683	-0.13
Dru_10745_42	A/G	0.31	0.69	0.31	0.54	0.44	0.359	-0.26
Dru_11110_29	A/T	0.48	0.52	0.48	0.46	0.51	0.695	0.08
Dru_11663_70	C/T	0.92	0.08	0.08	0.17	0.16	1.000	-0.09
Dru_11665_59	A/C	0.02	0.98	0.02	0.04	0.04	1.000	-0.02
Dru_11991_42	A/C	0.21	0.79	0.21	0.33	0.34	1.000	-0.01
Dru_12114_80	A/G	0.10	0.90	0.10	0.21	0.19	1.000	-0.12
Dru_12360_79	C/G	0.83	0.17	0.17	0.25	0.28	0.500	0.10
Dru_12407_94	A/C	0.67	0.33	0.33	0.33	0.45	0.350	0.25
Dru_12984_58	C/T	0.00	1.00	0.00	0.00	*	*	*
Dru_13283_60	A/T	0.10	0.90	0.10	0.21	0.19	1.000	-0.12
Dru_13374_84	A/G	0.33	0.67	0.33	0.42	0.45	1.000	0.06
Dru_13606_75	C/T	0.88	0.13	0.13	0.17	0.22	0.297	0.24
Dru_13817_38	A/T	0.98	0.02	0.02	0.04	0.04	1.000	-0.02
Dru_13823_79	A/T	0.96	0.04	0.04	0.08	0.08	1.000	-0.04
Dru_14305_38	A/G	0.17	0.83	0.17	0.33	0.28	1.000	-0.20
Dru_14440_51	C/T	0.98	0.02	0.02	0.04	0.04	1.000	-0.02
Dru_14682_52	A/G	0.75	0.25	0.25	0.42	0.38	1.000	-0.11
Dru_14684_61	C/T	0.19	0.81	0.19	0.38	0.31	0.550	-0.23
Dru_15113_62	A/G	0.92	0.08	0.08	0.08	0.16	0.126	0.45
Dru_15492_64	C/T	0.21	0.79	0.21	0.42	0.34	0.539	-0.26
Dru_16836_71	C/T	0.13	0.88	0.13	0.25	0.22	1.000	-0.14
Dru_17261_42	A/T	0.04	0.96	0.04	0.08	0.08	1.000	-0.04
Dru_17482_65	A/G	0.21	0.79	0.21	0.25	0.34	0.232	0.24
Dru_17913_48	A/C	0.29	0.71	0.29	0.33	0.42	0.347	0.19
Dru_18067_44	A/G	0.83	0.17	0.17	0.17	0.28	0.090	0.40
Dru_18398_48	C/T	0.71	0.29	0.29	0.42	0.42	1.000	-0.01

Dru_18730_45	C/G	0.90	0.10	0.10	0.21	0.19	1.000	-0.12
Dru_18801_70	A/T	0.48	0.52	0.48	0.46	0.51	0.694	0.08
Dru_19498_36	A/G	0.19	0.81	0.19	0.38	0.31	0.551	-0.23
Dru_19630_73	C/T	0.44	0.56	0.44	0.38	0.50	0.240	0.24
Dru_19751_33	C/T	0.25	0.75	0.25	0.42	0.38	1.000	-0.11
Dru_19790_29	C/G	0.88	0.13	0.13	0.25	0.22	1.000	-0.14
Dru_20186_51	A/G	0.23	0.77	0.23	0.21	0.36	0.062	0.41
Dru_20729_33	C/T	0.67	0.33	0.33	0.50	0.45	1.000	-0.13
Dru_21056_30	A/G	0.23	0.77	0.23	0.29	0.36	0.556	0.17
Dru_21066_50	A/G	0.46	0.54	0.46	0.67	0.51	0.213	-0.34
Dru_21326_30	A/G	0.48	0.52	0.48	0.46	0.51	0.696	0.08
Dru_21353_80	A/G	0.83	0.17	0.17	0.33	0.28	1.000	-0.20
Dru_23429_73	C/T	0.10	0.90	0.10	0.21	0.19	1.000	-0.12
Dru_23680_83	C/T	0.48	0.52	0.48	0.46	0.51	0.695	0.08
Dru_24606_73	A/T	0.31	0.69	0.31	0.38	0.44	0.636	0.13
Dru_26588_41	A/G	0.90	0.10	0.10	0.21	0.19	1.000	-0.12
Dru_26852_67	A/G	0.52	0.48	0.48	0.54	0.51	1.000	-0.09
Dru_27097_36	A/T	0.02	0.98	0.02	0.04	0.04	1.000	-0.02
Dru_27363_49	C/T	0.25	0.75	0.25	0.42	0.38	1.000	-0.11
Dru_29127_70	A/T	0.67	0.33	0.33	0.50	0.45	1.000	-0.13
Dru_29307_74	C/T	1.00	0.00	0.00	0.00	*	*	*
Dru_29342_37	G/T	0.10	0.90	0.10	0.21	0.19	1.000	-0.12
Dru_29503_83	A/G	0.92	0.08	0.08	0.17	0.16	1.000	-0.09
Dru_29746_68	C/G	0.71	0.29	0.29	0.58	0.42	0.126	-0.41
Dru_30966_28	C/T	0.75	0.25	0.25	0.33	0.38	0.596	0.11
Dru_31092_41	C/T	0.23	0.77	0.23	0.38	0.36	1.000	-0.06
Dru_33223_30	C/G	0.27	0.73	0.27	0.46	0.40	0.636	-0.16

Dru_34000_83	C/T	0.85	0.15	0.15	0.29	0.25	1.000	-0.17
Dru_34029_70	A/G	0.79	0.21	0.21	0.17	0.34	0.032	0.49
Dru_34145_54	C/T	0.10	0.90	0.10	0.21	0.19	1.000	-0.12
Dru_34390_25	A/C	0.88	0.13	0.13	0.25	0.22	1.000	-0.14
Dru_35116_58	A/G	0.81	0.19	0.19	0.38	0.31	0.551	-0.23
Dru_35252_83	A/G	0.54	0.46	0.46	0.42	0.51	0.433	0.16
Dru_35376_86	A/C	0.06	0.94	0.06	0.13	0.12	1.000	-0.07
Dru_36084_39	G/T	0.10	0.90	0.10	0.13	0.19	0.206	0.33
Dru_36629_70	A/T	0.00	1.00	0.00	0.00	*	*	*
Dru_36680_36	C/T	0.44	0.56	0.44	0.54	0.50	1.000	-0.10
Dru_37928_79	A/T	0.06	0.94	0.06	0.04	0.12	0.063	0.64
Dru_37990_83	C/T	0.19	0.81	0.19	0.38	0.31	0.551	-0.23
Dru_38472_27	A/T	0.63	0.38	0.38	0.58	0.48	0.391	-0.24
Dru_38632_81	C/T	0.13	0.88	0.13	0.25	0.22	1.000	-0.14
Dru_54499_36	C/T	0.52	0.48	0.48	0.46	0.51	0.695	0.08
Dru_55509_70	A/T	0.73	0.27	0.27	0.46	0.40	0.635	-0.16
Dru_61113_60	A/T	0.17	0.83	0.17	0.25	0.28	0.501	0.10
Dru_63639_60	A/T	0.48	0.52	0.48	0.63	0.51	0.410	-0.25
Dru_73417_43	A/G	0.50	0.50	0.50	0.50	0.51	1.000	0.00
Dru_74279_55	A/C	0.56	0.44	0.44	0.38	0.50	0.240	0.24

Minor allele frequency (MAF); observed heterozygosity (H_O); expected heterozygosity (H_E); probability of deviation from Hardy–Weinberg equilibrium (P_{HWE}); fixation index (F); not calculated as the SNP was monomorphic (*)